Journal of
Synchrotron
Radiation

ISSN 1600-5775

Received 17 May 2014
Accepted 16 July 2014

High-throughput synchrotron X-ray diffraction for combinatorial phase mapping

J. M. Gregoire,[a]* D. G. Van Campen,[b] C. E. Miller,[b] R. J. R. Jones,[a] S. K. Suram[a] and A. Mehta[b]*

[a]Joint Center for Artificial Photosynthesis, California Institute of Technology, Pasadena, CA 91125, USA, and [b]Stanford Linear Accelerator Laboratory, Stanford University, Menlo Park, CA 94025, USA. *E-mail: gregoire@caltech.edu, mehta@slac.stanford.edu

Discovery of new materials drives the deployment of new technologies. Complex technological requirements demand precisely tailored material functionalities, and materials scientists are driven to search for these new materials in compositionally complex and often non-equilibrium spaces containing three, four or more elements. The phase behavior of these high-order composition spaces is mostly unknown and unexplored. High-throughput methods can offer strategies for efficiently searching complex and multi-dimensional material genomes for these much needed new materials and can also suggest a processing pathway for synthesizing them. However, high-throughput structural characterization is still relatively under-developed for rapid material discovery. Here, a synchrotron X-ray diffraction and fluorescence experiment for rapid measurement of both X-ray powder patterns and compositions for an array of samples in a material library is presented. The experiment is capable of measuring more than 5000 samples per day, as demonstrated by the acquisition of high-quality powder patterns in a bismuth–vanadium–iron oxide composition library. A detailed discussion of the scattering geometry and its ability to be tailored for different material systems is provided, with specific attention given to the characterization of fiber textured thin films. The described prototype facility is capable of meeting the structural characterization needs for the first generation of high-throughput material genomic searches.

Keywords: combinatorial materials science; high-throughput phase mapping; X-ray diffraction; X-ray fluorescence.

1. Introduction

High-throughput (HiTp) methods offer opportunities for accelerated discovery of new functional materials, and can additionally enable the investigation of new structure–property relationships (Green *et al.*, 2013; Potyrailo *et al.*, 2011). HiTp computational investigation of an inorganic material typically involves the creation of an energetically relaxed single-crystal model and subsequent computation of its properties (Curtarolo *et al.*, 2013; Jain *et al.*, 2013). With this approach, the structure and property information are both readily available for investigation of structure–property relationships. HiTp experimental investigations typically involve the synthesis of a library of materials with varying composition, synthesis temperature or other processing conditions (Green *et al.*, 2013; Potyrailo *et al.*, 2011). Automated measurements of the material properties of interest are then performed, typically involving a measurement of performance for a target application. For a deeper understanding of the underlying materials science of the observed composition–processing–property trends, HiTp structural characterization through measurement of the crystalline phase(s) of each sample in the library is necessary. Phase mapping of material libraries has profound value for data informatics and provides a platform for comparing experimental and theoretic materials properties, a missing link in the development of strategies for materials by design. We have developed a synchrotron-based instrument for performing these phase mapping experiments on a wide variety of material library formats.

For inorganic materials, powder X-ray diffraction (XRD) is a ubiquitous structure characterization technique that can be paired with X-ray fluorescence (XRF) measurements to provide simultaneous composition characterization. A few combinatorial XRD/XRF techniques have been developed at synchrotron radiation facilities to accelerate data acquisition by exploiting the high monochromatic photon flux. Isaacs *et al.* (1998) used a focused X-ray microbeam to provide high spatial resolution in an experimental setup that afforded use

of several traditional techniques, including traditional XRF and XRD experiments with a point detector. Vogt *et al.* (2004) emphasized development of combinatorial XRF measurements in conjunction with a traditional (low-throughput) diffraction setup. While these early experiments included a diffractometer and point detector, more recent developments have employed area detectors, including a combinatorial diffraction experiment for characterizing epitaxital thin films (Kukuruznyak *et al.*, 2007). The high X-ray energy and transmission geometry of this experiment provide efficient detection of a wide region of reciprocal space. This configuration has been developed for efficient characterization of material libraries exhibiting fiber texture (Gregoire *et al.*, 2009) and also adapted to provide simultaneous acquisition of calorimetry data (Gregoire *et al.*, 2013). Transmission XRD on a substrate-supported material library creates inherent experimental limitations and strong background signal resulting from the scattering of the substrate. We present a new HiTp XRD/XRF experiment developed at the Stanford Synchrotron Radiation Lightsource (SSRL) at SLAC National Accelerator Laboratory with careful optimization of the scattering geometry for thin film samples. Through scattering simulation, the geometry of the material library and diffraction detector with respect to the synchrotron X-ray beam are optimized for a specific application, and we discuss the inherent compromises and provide guidance on optimization of the geometry for different applications. Simultaneous with the XRD data acquisition, an energy-resolving X-ray detector provides X-ray fluorescence characterization of sample composition.

In the fields of combinatorial and HiTp materials science, a wide variety of sample formats have been reported (Green *et al.*, 2013; Potyrailo *et al.*, 2011). While we have designed the HiTp experiment to be readily adapted to various formats, for the purposes of this paper we describe the experiment in the context of a thin film material library on a planar substrate. The set of samples to be measured are at a known set of positions on the library plate, and the HiTp experiment proceeds by performing a scattering/fluorescence experiment at each library position. The thickness of the thin film samples is not a critical parameter for the design of the experiment but is considered to be of the order of 1 μm or thinner. The lateral dimension of each sample is much more critical and the importance of this dimension is noted at several points in the paper. The presented scattering geometry was designed for a 1 mm × 1 mm sample size, which is well matched to the typical dimension in reported library synthesis techniques (Green *et al.*, 2013). The library substrate specifications are also considered in several experiment design aspects, and a 100 mm-diameter Si wafer will be used as an illustrative example.

2. Experimental

2.1. Scattering geometry

In designing the scattering geometry for a HiTp XRD experiment, consideration must be given not only to the quality of the XRD data but also to the ability to automate both the data acquisition and data processing. The scattering geometry must yield XRD powder patterns (intensity as a function of the magnitude of the scattering vector **Q**), and two-dimensional XRD patterns (intensity as a function of the magnitude and direction of the scattering vector **Q**) for textured materials at high throughput with moderate resolution of approximately 0.1 nm^{-1}. The use of an area detector to capture substantial portions of the scattering Debye rings can provide substantial peak intensity in powder patterns even with low integration time. For samples that exhibit fiber texture, area detectors play an increasingly important role of simultaneously measuring scattered intensity as a function of both **Q** and Ψ, the fiber texture angle (see Fig. 1). While a detailed description of fiber texture is beyond the scope of this paper, we note that vapor-deposited thin films often exhibit fiber texture in which a particular crystallographic direction ⟨*hkl*⟩ (where *h*, *k*, *l* are Miller indices) is aligned with substrate normal (**n**) (Ohring, 2001). The texture is manifested in scattering experiments by a non-uniform scattering intensity along the Debye ring of a given {*hkl*}. In the extreme case, finite scattering intensity is only observed at discrete values of the angle Ψ between **n** and **Q** and, if the scattering geometry does

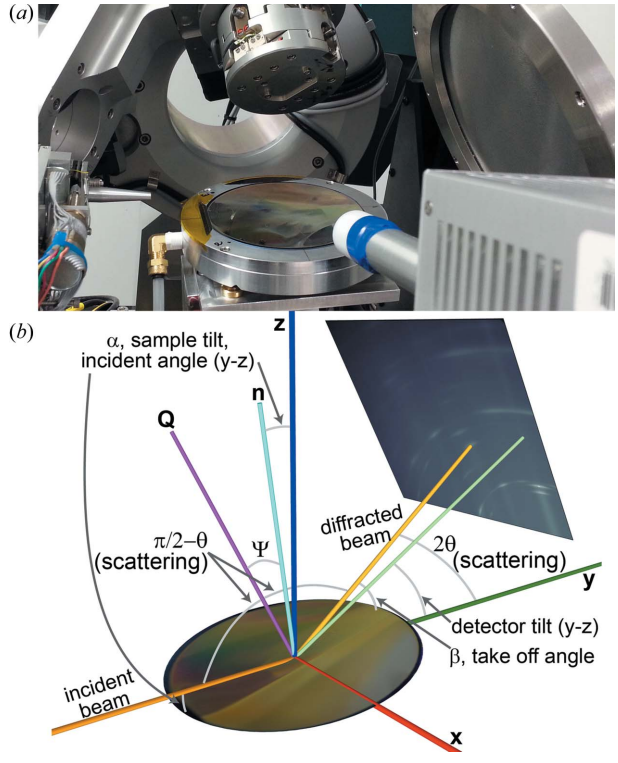

Figure 1
(*a*) Photograph of the combined XRD/XRF experiment on a continuous composition spread thin film library. (*b*) Illustration of the scattering geometry showing the important parameters for the material library plate (photograph of continuous composition spread thin film) and detector (example diffraction image). The geometry parameters for a particular scattered beam are also shown. A Cartesian coordinate system is noted and for each angle the parenthesis label denotes whether the angle lies in the scattering plane or a Cartesian axis plane. The angles Ψ and β do not lie in any of these planes.

not provide detection of the corresponding region of reciprocal space, the resulting powder pattern will be 'missing' peaks. Coverage of a 90° range of Ψ is sufficient for detection of at least one (*hkl*) for every {*hkl*} in cubic (and many other) structures, and we assert that a 60° range is sufficient for commonly observed combinations of crystal structure and texture direction. Depending on the scattering geometry, a Debye ring can include anywhere between a single value of Ψ and a wide range of angles. To provide robust characterization of textured films, we design the HiTp scattering geometry to cover a wide range of both **Q** and Ψ.

For maximum reciprocal space coverage in a single diffraction image, the use of high X-ray energy (30–100 keV) in transmission geometry is a very powerful technique (Kukuruznyak *et al.*, 2007; Gregoire *et al.*, 2009). The high X-ray energy condenses Debye rings of interest to small scattering angles, allowing a flat area detector to cover a large range of Ψ for every measured **Q**. This scattering geometry relies upon sufficient transparency of the substrate and even then the scattering signal is often dominated by the substrate signal, and small variation in the scattering from the substrate introduces artifacts that hamper automated data-processing routines. To reduce the scattering from substrate and to accommodate thicker (less X-ray transparent) substrate materials, we employ a reflection scattering geometry, which requires lower X-ray energy and careful design of the scattering geometry.

To describe the scattering geometry, we will consider a real-space coordinate system in which the X-ray beam travels in the **y** direction and the plane of the library plate is offset from the **x**–**y** plane by rotation about **x**, such that substrate normal **n** is displaced by an angle α from **z** (see Fig. 1). This angle α is the angle of incidence of the X-ray beam on the library plane. To minimize contribution to the XRD signal from the substrate and to provide the largest reciprocal space coverage, α should be slightly larger than the critical angle for total external reflection (θ_{crit}) of the incidence X-rays (<1°). However, working at such shallow incidence angles raises experimental challenges. Since θ_{crit} depends on density as well as chemistry, the value will change across a composition library. These variations θ_{crit} and small variations in α due to slight misalignment can yield substantial modulations in the diffraction intensity for a material library, which is undesirable for these combinatorial measurements. Furthermore, the vertical dimension of the beam spreads by a factor of $\sin(\alpha)^{-1}$ onto the sample, requiring a very narrow beam for $\alpha < 1°$. The incident X-ray beam cannot be narrowed by aggressive focusing as a scattered beam would be highly divergent, yielding poor **Q** resolution. Hence the incidence beam can be made narrow only by severe slitting, and the resultant loss in flux would be unacceptable for measurement throughput. A judicious balance between magnitude of the contribution from the substrate, misalignment and other artifacts and challenges arising from library flatness and non-uniformity must be reached. After several trials with varied X-ray energy, incidence angle and beam height, we used $\alpha = 6°$ with 1 mm × 0.1 mm beam for the measurements presented below. The corresponding beam spreading factor is approximately 10 to attain the desired 1 mm × 1 mm footprint on the sample, and this configuration was found to be optimal from the standpoints of data quality and practical operation of the HiTp experiment.

The detector used in this study is a Princeton Quad-RO 4320 with fiber optic couple CCD chip, providing effective imaging of a 120 mm × 120 mm square. To achieve fast readout time the detector output was binned 2 × 2 to give diffraction patterns with 1084 × 1084 pixels. With the X-ray beam and library geometry discussed above, and the given pixel size and area of the detector, its placement had to be optimized to obtain largest reciprocal space coverage with sufficient **Q** resolution. In the experiments described herein, the detector plane was tilted by 30° about **x** from the **x**–**z** plane, as shown in Fig. 1. Centering the detector along **x** provides duplicate information on each half of the detector due to equal **Q**–Ψ coverage. To expand the Ψ coverage, the detector was shifted by 50 mm along **x**, the detector horizontal. The detector was also shifted along in $-\mathbf{z}$ by 7 mm to further maximize reciprocal space coverage. Finally, the distance from the detector plane to the sample (light green line in Fig. 1) was 120 mm.

With the detector in place, we have described the complete scattering geometry depicted in Fig. 1 and will ascertain the scattering parameters over the array of detector pixels. While this geometry was designed for a particular X-ray energy, it is worth noting that the detector coverage of the three primary angles of interest is independent of X-ray energy. In Fig. 2 these angles mapped over the detector range, including the diffraction angle (2θ), the take-off angle between the diffracted beam and the library plane (β), and the fiber texture angle (Ψ).

This configuration pushes the limit of reflection geometry as the values of β reach values as low as 2° on the bottom edge of the detector. At this low angle, diffracted X-rays are attenuated by the sample and neighboring material on the library plate. Lowering the detector to this limit is important as it provides access to low-**Q**, expands Ψ-coverage (see Fig. 2) and helps mitigate defocusing. For X-rays scattered by a particular **Q**, the diameter of the diffracted beam is proportional to $\sin(\beta)$; the associated defocusing is negligible on the bottom of the detector and near the top reaches a value of 0.7 mm for the 1 mm beam projection on the sample. The loss of **Q**-resolution due to this defocusing depends on X-ray energy but is substantial and weakens high-**Q** peak intensity in the measured powder patterns.

2.2. X-ray energy

With the presented configuration, X-ray energy is chosen to provide the desired range and resolution of **Q** under consideration of the β and Ψ maps of Fig. 2. For the purpose of HiTp XRD characterization of unknown materials, the specific **Q** range of interest is that necessary for initial phase identification and evaluation of the composition–phase relationship. Because even a high-symmetry crystal structure has several

often closely spaced and sometimes overlapping powder diffraction peaks, the first few diffraction peaks, at the lowest **Q**, are of inordinate importance in identifying the structural phase of an unknown material. We note that the **Q** range needed for phase identification and discrimination may not be sufficient for a detailed structural refinement. The HiTp experiment described here involves characterization of metal oxide libraries in which known phases have somewhat large unit cells, yielding powder diffraction peaks below 20 nm^{-1} and occasionally approaching 10 nm^{-1}. The X-ray energy must be no greater than 13.5 keV to reach these values with the presented geometry. While various X-ray energies below this value can be chosen to provide high resolution or to limit X-ray penetration depth, we use this 13.5 keV X-ray energy to demonstrate wide-**Q** coverage in a single diffraction image. The corresponding map of **Q** is shown in Fig. 2.

Fig. 2 shows that Ψ-coverage is in excess of 60° for the middle range of **Q** values, with smaller Ψ-coverage at the extreme **Q** values. The highest **Q** values are only reached in the upper-left corner of the detector. For the present purposes, we ignore this region and perform analysis over a **Q** range of 9.3 to 56 nm^{-1}, which generally contains sufficient information for evaluation of phase behavior of inorganic systems.

The bending-magnet beamline 1-5 at SSRL, used for these measurements, has a Si 111 double-crystal monochromator followed by water-cooled Si toroidal mirror. The resulting flux at 13.5 keV in the 1 mm × 0.1 mm beam was approximately 5×10^{10} photon s^{-1}.

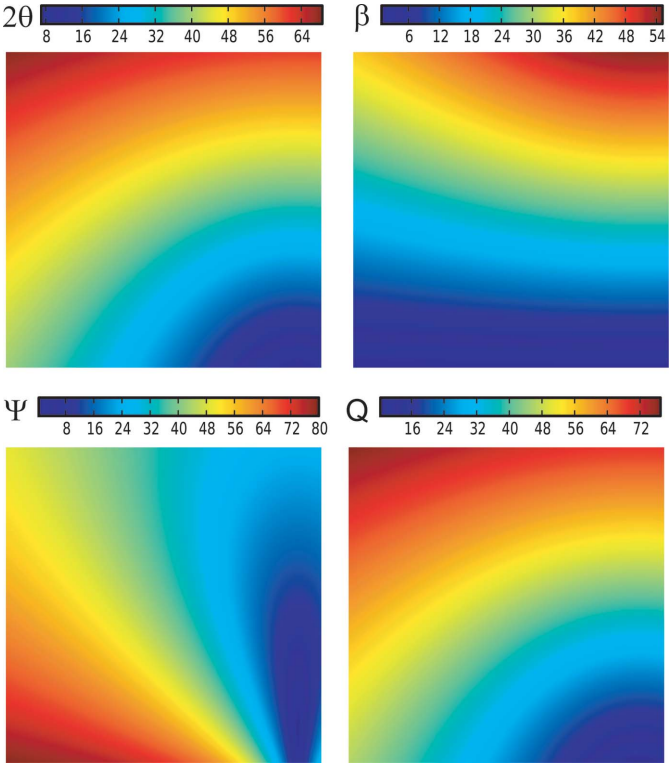

Figure 2
For the geometry shown in Fig. 1, the values of the three angles describing the scattering geometry and the scattering vector magnitude are mapped over the extents of the 120 mm square detector.

2.3. Library positioning

The basic positioning requirement for an automated experiment is a lateral translation stage to effect rastering of the X-ray beam across the sample library. Translation stages with range and resolution to match typical library formats are readily available, and we used a pair of precision Parker linear stages. These stages were mounted on a Huber rotational arc with the rotation axis nominally perpendicular to the X-ray beam for controlling the tilt angle α. The translation stages are accessible by the data acquisition computer, which coordinates the diffraction images sequence with library position. The more nuanced aspects of library position are in the control of sample flatness and levelness to ensure that the translation stage coordinates and library coordinates are commensurate. The detector was mounted on a large custom-built positioning gantry with four degrees of motion, three translations and one tilt. One of the translation motions controlled the sample-to-detector distance, whereas the other two moved it in the transverse directions.

The low angle of X-ray beam incidence on the plane of the library places significant constraints on the flatness of the samples and the library plate. For $\alpha = 6°$, *i.e.* a beam spreading factor of 10, variation δz in the sample height of 10 µm will displace the beam by 100 µm. This displacement corresponds to 10% of the beam footprint or library sample, which is an undesirably large misalignment. For the thin film application, sample thickness and thus sample roughness are much smaller than this 10 µm length scale, and thus the primary concern is the flatness of the substrate. It is worth noting that the flatness specification on many low-cost substrates such as glass and metal plates exceeds this flatness requirement, but we will continue to use Si wafers as an illustrative example. Various grades of Si wafers meet the 10 µm flatness specification, but a wafer does not necessarily remain flat upon deposition of a sample library. Consider a 100 mm-diameter 0.5 mm-thick Si substrate coated with a continuous thin film composition spread that is 500 nm thick with uniform internal stress. The thin film stress induces curvature in the substrate, and when resting on a flat surface the curvature introduces a change in substrate height between the center and edge of the substrate. A relatively low film stress of 120 MPa induces a radius of curvature of 125 m, corresponding to δz = 10 µm. A larger common stress of 1 GPa yields 15 m radius of curvature and δz = 83 µm, well above the acceptable level.

To mitigate alignment artifacts due to curvature in flexible substrates, the translation stage was outfit with a specially designed vacuum chuck. The vacuum chuck contained an array of 392 1.25 mm holes distributed over a 90 mm-diameter circle. After drilling the array of holes, the mounting surface of the chuck was lapped to a flatness specification of 1 µm to provide a planar surface for the substrate. Metrology measurements indicated that each hole is small enough to avoid local deformation of the substrate but the uniform

suction is sufficient to flatten any initial substrate curvature for a thin film library on 100 mm Si wafer. While the 10 μm flatness specification is to prevent spurious alignment artifacts, a related specification must be applied to the levelness of the substrate plane with respect to the translation plane. These planes must be level to within 0.1 mrad (10 μm elevation over 100 mm library position) to avoid systematic discrepancy between the translation stage coordinates and library coordinates. To attain fine leveling control, the vacuum chuck was mounted on a three-point leveling system, and the elevation of each point with respect to the translation stage was controlled by a Newfocus piezo actuator with incremental motion sensitivity of better than 30 nm. This system provides sufficient control for leveling the library plane but does not provide a measure of the levelness. Several measurement strategies involving the use of the incident beam could be employed, but we introduced an additional alignment tool. An Omron ZX-LD30V laser range finder was mounted above the nominal location for beam incidence. The laser range finder allowed us to ensure that the height of the combinatorial library moved by less than 1 μm with respect to the diffractometer and the beam-defining slits as the positioning system was rastered over the relevant range. It is worth noting that, with the addition of a z-positioning stage underneath the translation stages, the range finder can be used in active feedback to either level the library plate on the fly or correct the z-position of samples due to library flatness variation. Such an automatic leveling mechanism will be implemented in a subsequent revision of this experiment, which will allow a wide variety of substrates.

2.4. X-ray fluorescence

To characterize the sample composition during each diffraction experiment, a single-element Si drift detector was placed in the **x**–**z** plane near the +**x** axis (see Fig. 1). The detector was displaced by approximately 300 mm from the sample and fit with a W collimator with sub-1 mm pinhole to eliminate stray elastic scatter. The detector output was processed by a XIA digital signal processor, and converted to an XRF energy spectrum with 10 eV/channel resolution for every sample in the material library. Quantitative compositional analysis of thin film libraries using such XRF spectra has been shown in previous reports, including automated computation of the film composition and thickness from each XRF spectrum (Vogt *et al.*, 2004; Gregoire *et al.*, 2010). For the present purposes, we select fluorescence peaks of interest and create single-channel analyzers (SCAs) for the corresponding energy ranges. The total fluorescence intensity for each SCA is then measured during each XRD measurement.

3. Results and discussion

To demonstrate the performance of the XRD/XRF experiment, we present results from one single uniform composition thin film and one multi-element combinatorial composition library. Fig. 3 shows diffraction images from a uniform 200 nm-thick Hf film deposited on a 100 mm Si *via* magnetron sputtering. The highly textured thin film creates an elaborate pattern in which the intensity along a constant-**Q** arc (see Fig. 2) varies dramatically. Fig. 3(*a*) is the result of a 1 s exposure demonstrating that the scattering from the film is well characterized by this relatively short exposure. Fig. 3(*b*) is the result of a 30 s exposure in which some additional low-intensity features are captured. However, the detector regions with most intense scattering have saturated. These images can be used collectively to study the texture in this film, but this effort is beyond the scope of the present report.

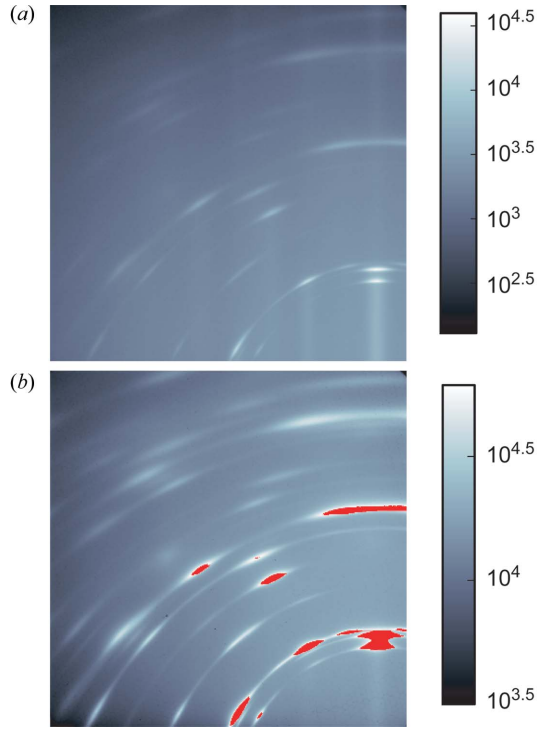

Figure 3
Diffraction images from a textured Hf film are shown with (*a*) 1 s and (*b*) 30 s exposure. The red regions denote saturation of the detector.

To demonstrate the combinatorial screening capability of the experiment, 177 XRD/XRF measurements were made on a three-cation oxide composition library. The continuous composition spread of Bi, V and Fe was synthesized by magnetron co-sputtering from elemental sources arranged symmetrically with respect to the 100 mm Si/SiO_2 substrate. The spatial variation in deposition rate from each source resulted in a continuous thin film with composition gradient of the order of 1 at.% mm^{-1}. The film thickness varied across the library within a factor of two of the center thickness of 200 nm. The relative deposition rates were chosen to provide a composition of $Bi_{40}Fe_5V_{55}$ at the substrate center. The metal composition library was calcined in air at 773 K, producing the metal oxide library used in the XRD/XRF experiment. Measurements were taken on a regular grid with 6 mm pitch and within a radius of 44 mm from the substrate center.

Fig. 4(*a*) shows a 100 s XRF spectrum at the center of the library. The highlighted regions indicate the selected windows for the three SCA channels for the elements of interest. For each of the library locations shown in Fig. 4(*b*), 15 s XRD/

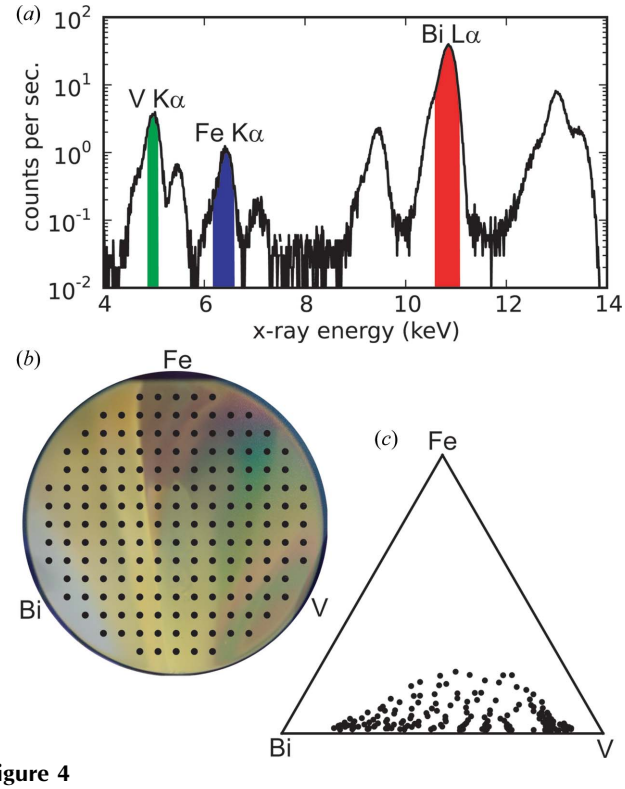

Figure 4
(*a*) XRF spectrum collected at the center of the continuous composition spread shown in (*b*). The fluorescence intensities in the three regions noted in (*a*) were measured at the 177 positions shown in (*b*) to determine the sample compositions, which are shown in (*c*).

XRF measurements were performed. Barring strong matrix effects, the assumption was made of proportionality of each atomic concentration to its respective SCA intensity. The XRF measurements of composition are represented in the ternary composition plot of Fig. 4(*c*).

A thin film of finely ground LaB_6 was placed on a similar Si wafer, and a diffraction pattern was collected to be used for subsequent calibration of the detector placement. The sample-to-detector distance and the detector tilt were calculated from the diffraction pattern from LaB_6. The average dark signal from the detector was subtracted from the raw images to produce diffraction images. Using the detector calibration from the LaB_6 pattern, 177 diffraction images were transformed into reciprocal space and the intensities along the azimuthal direction were integrated to produce powder patterns (intensity *versus* **Q**) using *WxDiff* (Mannsfeld *et al.*, 2011) software. To assess the quality of data resulting from this experiment, the 177 patterns are plotted in logarithmic scale (to enhance weak peaks of importance for phase identification) without any background subtraction, smoothing or other processing. The resulting 177 powder patterns are shown in Fig. 5(*a*), and the series of 15 patterns from the substrate horizontal are shown in Fig. 5(*b*). The phase behavior of pseudoternary oxide composition spaces can be quite complex, but visual inspection of Fig. 5(*b*) (especially the presence or absence of weak peaks) suggests the presence of five unique phases in this particular one-dimensional cut through the composition space.

Visual identification of five distinct phases from 15 diffraction patterns shown in Fig. 5(*b*) takes effort and significantly longer time than it took to collect the data. This task becomes exponentially more difficult for larger, textured combinatorial libraries. Collectively, Fig. 5 illustrates the need for automated processing of the HiTp scattering data. The development of such algorithms is an active field of research (Ermon *et al.*, 2012; Le Bras *et al.*, 2011; Long *et al.*, 2009; Baumes *et al.*, 2009) and we note that HiTp generation of high-quality powder patterns is necessary for extensive testing of recently developed algorithms. While we present the data from a single library in this report, the initial run of this experiment produced data from several metal oxide libraries. To foster the rapid development of automated data processing and interpretation algorithms, the data are being made publicly available and coupled with visualization tools and metrics for evaluating phase diagrams calculated from the powder patterns (Le Bras *et al.*, 2014).

With the integration time of 15 s for each sample and the 2–3 s downtime between measurements (sample translation and software initialization), the entire Bi–V–Fe oxide library dataset was collected in less than 1 h. At this rate, approxi-

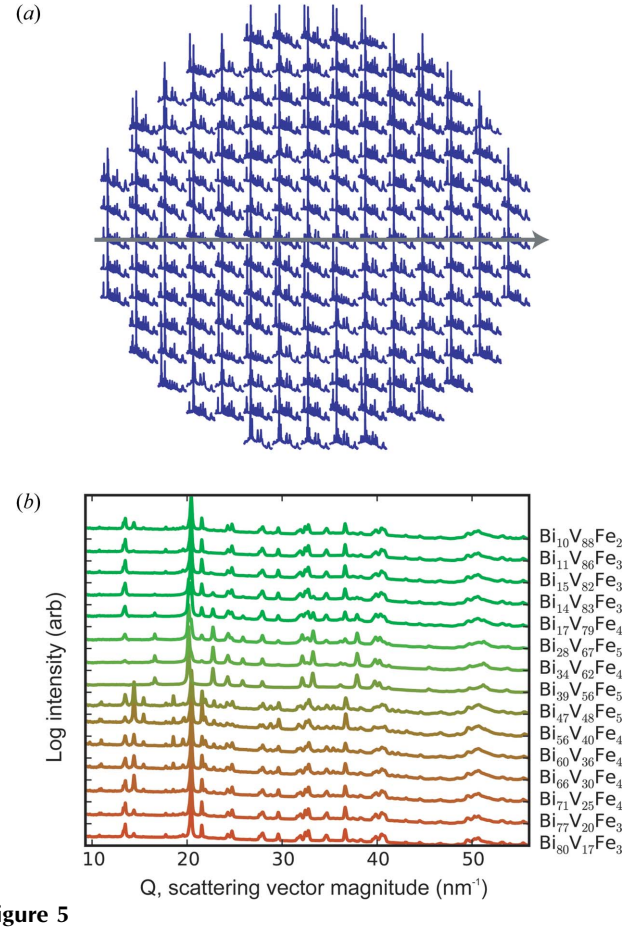

Figure 5
(*a*) The 177 powder patterns from a Bi–V–Fe oxide composition library are shown with log-scale intensity. Each pattern is shown at the corresponding library position (see Fig. 4*b*). (*b*) The patterns along the horizontal line shown in (*a*) are colored according to the measured compositions, which are listed to the right for each of the 15 patterns.

mately 5000 patterns can be collected per day. As shown in Fig. 3, some materials are amenable to structural characterization with 1 s integration time, demonstrating that this experiment can attain much higher throughput.

4. Conclusion

High-throughput experiments must be designed to not only provide the desired data but also facilitate robust operation with minimal sample-specific artifacts. Guided by these design principles, we developed a synchrotron experiment that enables phase mapping of composition libraries through the simultaneous collection of XRD and XRF data. The scattering geometry optimizes the competing interests of XRD reciprocal space coverage and practical experiment operation. Through both measurements and simulation of the reciprocal space coverage by the XRD area detector, we demonstrate efficient characterization of textured thin films with a single diffraction image per sample. Through rastering of a material library, high-quality powder patterns were collected with throughput of less than 18 s per sample, demonstrating the capacity for acquiring at least 5000 patterns per day. With this combination of throughput and data quality, the experiment is well suited for mapping phase behavior in high-order composition spaces, a crucial capability for developing structure–property relationships and connecting experimental observations with the computational materials databases.

This material is based upon work performed by the Joint Center for Artificial Photosynthesis, a DOE Energy Innovation Hub, supported through the Office of Science of the US Department of Energy (Award No. DE-SC0004993). Use of the Stanford Synchrotron Radiation Lightsource, SLAC National Accelerator Laboratory, is supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences under Contract No. DE-AC02-76SF00515. The authors thank Drs Matthew Kramer and Ichiro Takeuchi for insightful discussions on technique development. The authors also thank Drs Carla Gomes, Ronan Le Bras and Bruce van Dover for assistance with data processing..

References

Baumes, L. A., Moliner, M. & Corma, A. (2009). *Chem.Eur. J.* **15**, 4258–4269.

Curtarolo, S., Hart, G. L. W., Nardelli, M. B., Mingo, N., Sanvito, S. & Levy, O. (2013). *Nat. Mater.* **12**, 191–201.

Ermon, S., Le Bras, R., Gomes, C. P., Selman, B. & van Dover, R. B. (2012). *Proceedings of the 15th International Conference on Theory and Applications of Satisfiability Testing (SAT 2012)*, 17–20 June 2012 Trento, Italy, pp 172–185.

Green, M. L., Takeuchi, I. & Hattrick-Simpers, J. R. (2013). *J. Appl. Phys.* **113**, 231101.

Gregoire, J. M., Dale, D., Kazimirov, A., DiSalvo, F. J. & Bruce van Dover, R. (2010). *J. Vac. Sci. Technol. A*, **28**, 1279–1280.

Gregoire, J. M., Dale, D., Kazimirov, A., DiSalvo, F. J. & van Dover, R. B. (2009). *Rev. Sci. Instrum.* **80**, 123905.

Gregoire, J. M., Xiao, K., McCluskey, P. J., Dale, D., Cuddalorepatta, G. & Vlassak, J. J. (2013). *Appl. Phys. Lett.* **102**, 201902.

Isaacs, E. D., Marcus, M., Aeppli, G., Xiang, X.-D., Sun, X.-D., Schultz, P., Kao, H.-K., Cargill, G. S. & Haushalter, R. (1998). *Appl. Phys. Lett.* **73**, 1820.

Jain, A., Ong, S. P., Hautier, G., Chen, W., Richards, W. D., Dacek, S., Cholia, S., Gunter, D., Skinner, D., Ceder, G. & Persson, K. A. (2013). *Appl Phys. Lett. Mater.* **1**, 011002.

Kukuruznyak, D. A., Reichert, H., Okasinski, J., Dosch, H., Chikyow, T., Daniels, J. & Honkimäki, V. (2007). *Appl. Phys. Lett.* **91**, 071916.

Le Bras, R., Bernstein, R., Gregoire, J. M., Suram, S. K., Gomes, C. P., Selman, B. & van Dover, R. B. (2014). In *Proceedings of the Twenty-Eighth Conference on Artificial Intelligence (AAAI-14)*, Québec City, Canada.

Le Bras, R., Damoulas, T., Gregoire, J. M., Sabharwal, A., Gomes, C. P. & Dover, R. B. V. (2011). *Proceedings of the 17th International Conference on Principles and Practice of Constraint Programming*, pp. 508–522. Berlin: Springer.

Long, C. J., Bunker, D., Li, X., Karen, V. L. & Takeuchi, I. (2009). *Rev. Sci. Instrum.* **80**, 103902.

Mannsfeld, S. C. B., Tang, M. L. & Bao, Z. (2011). *Adv. Mater.* **23**, 127–131.

Ohring, M. (2001). *Materials Science of Thin Films.* Amsterdam: Elsevier Science.

Potyrailo, R., Rajan, K., Stoewe, K., Takeuchi, I., Chisholm, B. & Lam, H. (2011). *ACS Comb. Sci.* **13**, 579–633.

Vogt, S., Chu, Y. S., Tkachuk, A., Ilinski, P., Walko, D. A. & Tsui, F. (2004). *Appl. Surf. Sci.* **223**, 214–219.